



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/582,932	06/14/2006	Yoko Jacobson	30162/42153	8158
4743 7590 03/26/2008 MARSHALL, GERSTEIN & BORUN LLP 233 S. WACKER DRIVE, SUITE 6300 SEARS TOWER CHICAGO, IL 60606			EXAMINER BORSETTI, GREG	
			ART UNIT 4141	PAPER NUMBER
			MAIL DATE 03/26/2008	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/582,932

Applicant(s)

JACOBSON, YOKO

Examiner

GREG A. BORSETTI

Art Unit

4141

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 14 June 2006.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 24-42 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 24-42 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 14 June 2006 is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-8508)
Paper No(s)/Mail Date 6/14/2006
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Inventor's Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

1. Claims 24-42 are pending.
Claims 1-23 have been canceled

Information Disclosure Statement

2. The Information Disclosure Statement (IDS) submitted on 6/14/2006 is not in compliance with the provisions of 37 CFR 1.97.

Based upon MPEP 609 and § 1.56(c), examiner requests that English-language translations be provided for each of the foreign patent documents including

- JP-4-160473
- JP-2000-20524
- JP-2003-263434

Drawings

3. The drawings filed on 6/14/2006 are objected to by the examiner. Fig. 3 is not provided in English. Translation is needed.

Specification

4. The abstract of the disclosure is objected to because the abstract must be self-withstanding. There are to be no references (numbering) to refer to other parts of the specification. Furthermore, the abstract should be a concise summary of the invention not totaling over 150 words. Correction is required. See MPEP § 608.01(b).

The amended Claim 38 is not labeled (New, Canceled, Amended, etc.) and appears to be a duplicate of Claim 39. Correction is required.

Claim 40 ends with a comma. It must end with a period. Correction is required.

The specification is objected to for not disclosing the tables as mentioned within the specification. Tables 1-6 on pages 54, 56, and 59 of the amended clean version of the specification are not shown although they are referred to. Correction is needed.

The lengthy specification has not been checked to the extent necessary to determine the presence of all possible minor errors. Applicant's cooperation is requested in correcting any errors of which applicant may become aware in the specification.

The claims are generally narrative and indefinite, failing to conform with current U.S. practice. They appear to be a literal translation into English from a foreign document and are replete with grammatical and idiomatic errors. Correction is required.

- For example, ¶ [0100] states "An upstream word having the shortest distance from the phrase to be translated (this time, an upstream word adjacent to the phrase to be translated) is extracted from an upstream word group existing in the upstream side of the phrase to be translated in the original sentence, and a downstream word having the shortest distance from the phrase to be translated (this time, a downstream word adjacent to the phrase to be translated) is extracted from a downstream word group existing in the downstream side of the phrase to be translated in the original sentence. Then the position where the phrase to be translated exists is assumed to be the reference position in each

natural sentence containing the phrase to be translated extracted from the corresponding translation database by the retrieval (when the phrase to be translated exists in a plurality of places in the natural sentence, the position where any phrase to be translated exists is assumed to be the reference position) to search the upstream region within the number of specified words (for example within three words) from the reference position to the upstream side for the upstream word already extracted and to search the downstream region within the number of specified words (for example within three words) from the reference position to the downstream side for the downstream word already extracted." This paragraph is vague and ambiguous and the examiner does not fully understand the scope of the concept within the highlighted phrase. It should be understood that this is merely an example and that the specification is replete with vague language and should be corrected.

Claim Rejections - 35 USC § 112

5. Claims 28-32, 35-37, 41 and 42 rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The stated claims cite an "object phrase" where it is not specifically defined in the specification. For the purposes of examination, an object phrase is taken to be a phrase within a sentence that has specific meaning or value.

Claim 31 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The stated claim cites a storage means for storing, then a parallel translation identifying system consisting of units. It is not understood to which the claim is referring to. Correction is needed. For the purposes of examination, the parallel translation identifying system is assumed to be the system that the claim refers to.

Claims 24-42 are generally narrative and indefinite, failing to conform with current U.S. practice. They appear to be a literal translation into English from a foreign document and are replete with grammatical and idiomatic errors. It is requested by the examiner that the application be conformed to US standards and that the claim language be compiled to a more readable disclosure.

Claim Rejections - 35 USC § 101

6. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 24-42 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claims cite storage, retrieval, search, and identifying in which the language of the claims raise the question as to whether the claims are directed to an abstract idea that is not tied to a technological environment or machine. Appropriate output for tangibility is needed.

Claims 41-42 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claims cite a program that is not connected to a computer readable medium or tangible hardware device. The claims appear to be programs per se. Correction is required.

Claim Rejections - 35 USC § 102

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 28,35,40,42 are rejected under 35 U.S.C. 102(b) as being disclosed by Gastaldo et al. (Patent # 6473729).

As per claim 28, Gastaldo teaches:

- **a memorizing means storing multiple natural sentences composed of multiple words in a source language by matching the parallel sentences in a target language**
- [Gastaldo, column, lines] discloses "sentence database 17 stores a plurality of sentences in English as the source language, and also stores to each English sentence the corresponding sentence in French as the target language." A database teaches a memorizing means because a database stores incoming

data and furthermore Gastaldo teaches that there are corresponding sentences in a source and target language.

- **a search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in memory**
- [Gastaldo, column 5, lines 60-66] discloses "...**control unit 14 queries the phrase-indexed sentence database 17 in step 41**". Since all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, **the system is able to retrieve** all pairs of sentences indexed by the selected noun phrase without any substantive delay." Gastaldo discloses that retrieval of natural sentences corresponding to sentence pairs of original and translated sentences is performed for original sentences.
- **a parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the first evaluated value as at least a parallel equivalent for the object phrases to be rendered in the original text, respectively granting the first evaluated value concordant with the small distance in the original text from the object phrases to be rendered in the original text to the respective words in the original text, recognizing the word corresponding to the original text from among the respective words in the natural sentence in question regarding the natural sentence extracted through a process by the search**

unit, thus, computing by the natural sentence the total value of the first evaluated value granted to the identified concordant words

[Gastaldo, Claim 8] discloses "wherein said word phrase is extracted from a text fragment of said input text, and the step of obtaining a translation includes the step of sorting the revealed pairs of text fragments **according to the number of word phrases common with said text fragment of said input text.**" The translation takes into consideration the number of word phrases in common between the source and target, which is analogous to the first evaluated value.

Claim 35 is rejected under the same grounds as Claim 28 for having parallel limitations because Claim 28's system performs the method as described in Claim 35 and has been rejected above.

Claim 40 is rejected under the same grounds as Claim 28 for having parallel limitations because Claim 28's system performs the method for a computer program as described in Claim 40 and have been rejected above. Gastaldo discloses the use of a computer program in [Gastaldo, column 4, lines 35-38] "The interface software may for instance be written in Visual Basic, and the resulting application can be an OLE server which can be integrated directly to any Visual Basic or C/C++ code."

Claim 42 is rejected under the same grounds as Claim 28 for having parallel limitations because Claim 28's system performs the method for a computer program as described in Claim 42 and have been rejected above. Gastaldo discloses the use of a

computer program in [Gastaldo, column 4, lines 35-38] "The interface software may for instance be written in Visual Basic, and the resulting application can be an OLE server which can be integrated directly to any Visual Basic or C/C++ code."

Claim Rejections - 35 USC § 103

8. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 24,29,33,38,39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gastaldo et al. (Patent # 6473729) in view of Eisele (Pre-Grant Publication #20020107683).

As per claim 24, Gastaldo discloses:

- **a storage means for storing a plurality of natural sentences, of original sentences composed of a plurality of words, correlated with translated sentences in a target language**
- [Gastaldo, column 4, lines 47-50] discloses "Sentence database 17 stores a plurality of sentences in English as the source language, and also stores to each English sentence the corresponding sentence in French as the target

language." The database in Gastaldo is a well-known storage means which contains corresponding original and translated sentences.

- a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means

- [Gastaldo, column 5, lines 60-66] discloses "...control unit 14 queries the phrase-indexed sentence database 17 in step 41". Since all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, the system is able to retrieve all pairs of sentences indexed by the selected noun phrase without any substantive delay." Gastaldo discloses that retrieval of natural sentences corresponding to sentence pairs of original and translated sentences is performed for original sentences.

Gastaldo fails to teach,

- a search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing translations appearing in the same sentence of said source language
- a second translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and specific frequently appearing phrases for frequently appearing translations of phrase to be translated by referring to said specific frequently appearing phrases found by said search means and existing in

said original sentences and referring to translated sentences of natural sentences containing each of said phrase to be translated among natural sentences extracted by retrieval of said retrieval means and for identifying found frequently appearing translations as translations of said phrases to be translated in said original sentences

Eisele, in analogous art, teaches the above limitations,

- **a search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing translations appearing in the same sentence of said source language**
- [Eisele, ¶ 0033] discloses "First, the process performs monolingual pre-processing and extraction of monolingual lexical and structural information." [Eisele, Fig. 4] shows that monolingual pre-processing counts and stores frequencies of normalized forms for weighting. This teaches that phrases are counted for the purposes of determining higher translation probability for translated phrases.
- Eisele and Gastaldo are analogous art because they both teach translation method through comparison of bilingual corpora. It would be obvious to someone of ordinary skill in the art to combine Eisele with the Gastaldo device because [Eisele, ¶ 0007] discloses that "Techniques based on lexical information offer a high quality and more robustness, but at the price of increased computational complexity." Furthermore, [Eisele, ¶ 0008] discloses

"The present invention has being made in consideration of the above situation, and has as its primary object to provide a method and system for extracting translations from translated texts that offer high quality and are robust against noisy data, but which still run fast and reliably."

- **a second translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and specific frequently appearing phrases for frequently appearing translations of phrase to be translated by referring to said specific frequently appearing phrases found by said search means and existing in said original sentences and referring to translated sentences of natural sentences containing each of said phrase to be translated among natural sentences extracted by retrieval of said retrieval means and for identifying found frequently appearing translations as translations of said phrases to be translated in said original sentences**
- [Eisele, ¶ 0038] discloses "the list of potential translations is generated based on a similarity evaluation of the word forms and their frequencies. After generating a list of pairs of normalized forms in step 510, the pairs are sequentially accessed (steps 520, 590). If the forms are identical or almost identical and the frequencies of the forms and the respective documents are sufficiently similar, these forms are taken as translation candidates." A comparison of frequencies for both documents is performed and used to determine which of the phrases for the translated document are

applicable to the original phrase by comparing the frequencies of occurrence for the original and translated phrases.

- Eisele and Gastaldo are analogous art because they both teach translation method through comparison of bilingual corpora. It would be obvious to someone of ordinary skill in the art to combine Eisele with the Gastaldo device because [Eisele, ¶ 0007] discloses that "Techniques based on lexical information offer a high quality and more robustness, but at the price of increased computational complexity." Furthermore, [Eisele, ¶ 0008] discloses "The present invention has being made in consideration of the above situation, and has as its primary object to provide a method and system for extracting translations from translated texts that offer high quality and are robust against noisy data, but which still run fast and reliably."

As per claim 29, claim 28 is incorporated and Gastaldo fails to teach:

- **said translation identifying means identifies the parallel translation of the minimal phrases to be rendered in the parallel translation of the natural sentence selected based on the computed total value of the second evaluated value as the parallel translation for the minimal phrases in the original text, if there are multiple sentences from the computed result of the first evaluated value, about the respective natural sentences of such multiplicity, respectively assigning the second evaluated value according to the small distance in the above natural sentence from the object**

phrases for rendering in the natural sentence for the matching words in the said sentence, computing the total value of the second evaluated value given to the corresponding words aforementioned by the natural sentence

Eisele, in analogous art, teaches the above limitation,

- [Eisele, ¶ 0060] discloses "in this embodiment, a backward run of the Hunt/Szymanski algorithm is performed in advance and the intermediate results are recorded sequentially in a stack in such a way that they can be "replayed" in reverse order, i.e. during the left-to-right traversal of the alignment matches." The Hunt/Szymanski algorithm is a well-known method of obtaining an LCS (Longest Common Subsequence), which is well-known to be used in finding the difference between two strings using word order as a metric. This is analogous to the second evaluated value because the second evaluated value is a similarity of the arrangement of the ordering of the words between the original and natural sentences.
- Eisele and Gastaldo are analogous art because they both teach translation method through comparison of bilingual corpora. It would be obvious to someone of ordinary skill in the art to combine Eisele with the Gastaldo device because [Eisele, ¶ 0007] discloses that "Techniques based on lexical information offer a high quality and more robustness, but at the price of increased computational complexity." Furthermore, [Eisele, ¶ 0008] discloses "The present invention has being made in consideration of the above situation,

and has as its primary object to provide a method and system for extracting translations from translated texts that offer high quality and are robust against noisy data, but which still run fast and reliably.”

Claim 33 is rejected under the same grounds as Claim 24 for having parallel limitations because Claim 24's system performs the method as described in Claim 33 and has been rejected above.

Claims 38 and 39 are rejected under the same grounds as Claim 24 for having parallel limitations because Claim 24's system performs the method for a computer program as described in Claims 38 and 39 and have been rejected above. Gastaldo discloses the use of a computer program in [Gastaldo, column 4, lines 35-38] “The interface software may for instance be written in Visual Basic, and the resulting application can be an OLE server which can be integrated directly to any Visual Basic or C/C++ code.”

Claims 25,26,27,34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gastaldo et al. (Patent # 6473729) in view of Kozma (Patent # 5983169).

As per claim 25, Gastaldo teaches:

- A system for identifying a corresponding translation, comprising a storage means for storing a plurality of natural sentences in a source**

language composed of a plurality of words correlated with translated sentences in a target language

- [Gastaldo, column 4, lines 47-50] discloses "Sentence database 17 stores a plurality of sentences in English as the source language, and also stores to each English sentence the corresponding sentence in French as the target language." The database in Gastaldo is a well-known storage means which contains corresponding original and translated sentences.
- **a retrieval means for retrieving natural sentences containing phrases to be translated in original sentences in a source language from a plurality of natural sentences in a source language stored in said storage means**
- [Gastaldo, column 5, lines 60-66] discloses "...**control unit 14 queries the phrase-indexed sentence database 17 in step 41.** Since all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, **the system is able to retrieve** all pairs of sentences indexed by the selected noun phrase without any substantive delay." Gastaldo discloses that retrieval of natural sentences corresponding to sentence pairs of original and translated sentences is performed for original sentences.

Gastaldo fails to teach:

- **an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means**

- **a translation identifying means for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences**

Kozma, in analogous art, teaches the above limitations,

- **an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means**
- [Kozma, column 4, lines 34-43] discloses "In accordance with the present invention, the system in FIG. 2 further includes means 13 for recording multiple parse structures for single expressions from input 3 and for further recording structures of output expression candidates generated from such syntactic structures, means 15 for comparing the possible syntactic structures of expressions in input 3 to those of corresponding output expression candidates, and means 17 for transforming the structures recorded by means 13 into alternative structures with like meaning." Kozma teaches that alternative phrases are considered, which means they would have to be identified, while the meaning is not changed in order to reduce ambiguity in multi-lingual translation [Kozma, column 5, lines 65-67].

- Kozma and Gastaldo are analogous art because both pertain to machine translation of natural language documents. It would be obvious to combine Kozma with the Gastaldo device because a goal of Kozma "is to reduce the instances of errors in the output of machine translation systems due to structural misinterpretations of input expressions", which is taken into consideration in Gastaldo. [Kozma, column 2, lines 21-24]
- **a translation identifying means for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences**
- [Kozma, column, lines] discloses "The general procedure of the present invention as described above is to compare the possible structures of each input expression with those of its candidate output expressions to determine whether there is a candidate output expression whose possible interpretations match those of the input expression." Kozma compares the original phrase alternatives to the output translation of the translated phrases to correct for ambiguities in phrase structure prior to translation, which teaches the instant application.
- Kozma and Gastaldo are analogous art because both pertain to machine translation of natural language documents. It would be obvious to combine

Kozma with the Gastaldo device because a goal of Kozma "is to reduce the instances of errors in the output of machine translation systems due to structural misinterpretations of input expressions", which is taken into consideration in Gastaldo. [Kozma, column 2, lines 21-24]

As per claim 26, claim 25 is incorporated and Gastaldo teaches:

- **said identifying means searches a plurality of natural sentences stored in said storage means for natural sentences containing said targeted words**
- [Gastaldo, column, lines] discloses "turning now to FIGS. 3 and 4, which illustrate a flow chart of the translation process, the user inputs the text to be translated in step 31. The text is then displayed so that the user can select an unknown word or group of words, e.g. by double-clicking on the word. Once a word has been selected in step 32, **phrase extractor 15 extracts in step 33 all the possible noun phrases relating to the selected word.**" All phrases are searched that contain the targeted word where it would be obvious that by searching all phrases it would inherently search all sentences.

Gastaldo fails to teach:

- **searches natural sentences stored in said storage means for natural sentences having the same sentence structure as natural sentences extracted by said searching**
- **identifies phrases replaced with targeted phrases in natural sentences extracted by said retrieval as said alternative phrases.**

Kozma, in analogous art, teaches the above limitations,

- **searches natural sentences stored in said storage means for natural sentences having the same sentence structure as natural sentences extracted by said searching**
- [Kozma, column, lines] discloses "In accordance with the present invention, the possible interpretations or structures for each input expression are compared with those of each output expression candidate by comparing means 15, to determine whether a candidate exists whose interpretations match those of the input expression."
- Kozma and Gastaldo are analogous art because both pertain to machine translation of natural language documents. It would be obvious to combine Kozma with the Gastaldo device because a goal of Kozma "is to reduce the instances of errors in the output of machine translation systems due to structural misinterpretations of input expressions", which is taken into consideration in Gastaldo. [Kozma, column 2, lines 21-24]
- **identifies phrases replaced with targeted phrases in natural sentences extracted by said retrieval as said alternative phrases.**
- [Kozma, column 4, lines 34-43] discloses "In accordance with the present invention, the system in FIG. 2 further includes means 13 for recording multiple parse structures for single expressions from input 3 and for further recording structures of output expression candidates generated from such syntactic structures, means 15 for comparing the possible syntactic structures of

expressions in input 3 to those of corresponding output expression candidates, and means 17 for transforming the structures recorded by means 13 into alternative structures with like meaning." Kozma teaches that alternative phrases are considered, which means they would have to be identified, while the meaning is not changed in order to reduce ambiguity in multi-lingual translation [Kozma, column 5, lines 65-67].

- Kozma and Gastaldo are analogous art because both pertain to machine translation of natural language documents. It would be obvious to combine Kozma with the Gastaldo device because a goal of Kozma "is to reduce the instances of errors in the output of machine translation systems due to structural misinterpretations of input expressions", which is taken into consideration in Gastaldo. [Kozma, column 2, lines 21-24]

As per claim 27, claim 24 is incorporated and Gastaldo teaches:

- **said identifying means determines a degree of coincidence between the natural sentences extracted by the retrieval of the aforementioned retrieval means and the aforementioned original sentence**
- [Gastaldo, column 5, lines 60-66] discloses "Once the user has selected one of the possible noun phrases, control unit 14 queries the phrase-indexed sentence database 17 in step 41. Since **all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, the system is able to retrieve all pairs of sentences indexed by the selected noun**

phrase without any substantive delay. The system then sorts in step 42 the retrieved pairs of sentences according to their relatedness to the original sentence. If they retrieved pairs are sorted, there must be a degree of coincidence in order for there to be a bias for which to sort.

- **Identifies the translations of at least the aforementioned phrases to be translated in the translations of the natural sentences selected on the basis of the determined degree of coincidence as the translations of at least the aforementioned phrases to be translated in the aforementioned original sentence**
 - [Gastaldo, column 6, lines 13-17] discloses "Once the user has decided that the selected pair of translated sentences should be used for translating the selected word of the input text, **phrase extractor 15 extracts in step 45 from the selected sentence pair the noun phrase translation and inserts the translated noun phrase automatically into the translation of the input text.**"
- The translations of the phrases are selected on the basis of the determined degree of coincidence from the original sentence.

Claim 34 is rejected under the same grounds as Claim 25 for having parallel limitations because Claim 25's system performs the method as described in Claim 34 and has been rejected above.

Claims 30,31,36,37,41 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gastaldo et al. (Patent # 6473729) in view of Horiguchi et al. (Patent # 6243669).

As per claim 30, Gastaldo teaches:

- **a memorizing means storing multiple natural sentences composed of multiple words in source language by matching with the parallel sentences in target language**
- [Gastaldo, column, lines] discloses "sentence database 17 stores a plurality of sentences in English as the source language, and also stores to each English sentence the corresponding sentence in French as the target language." A database teaches a memorizing means because a database stores incoming data and furthermore Gastaldo teaches that there are corresponding sentences in a source and target language.
- **a search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in source language from among multiple natural sentences in source language stored in the memory**
- [Gastaldo, column 5, lines 60-66] discloses "...control unit 14 queries the phrase-indexed sentence database 17 in step 41. Since all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, the system is able to retrieve all pairs of sentences indexed by the selected noun phrase without any substantive delay." Gastaldo discloses that

retrieval of natural sentences corresponding to sentence pairs of original and translated sentences is performed for original sentences.

Gastaldo fails to teach,

- a parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the third evaluated value as at least a parallel equivalent for the object phrases to be rendered in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific anterior words when detected and the object phrases to be rendered or the anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the original text, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back

side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific posterior words when detected and the object phrases to be rendered or the posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the original text, and thus, computing by the natural sentence the total value of the 3rd evaluated value

Horiguchi, in analogous art, teaches the above limitations,

- [Horiguchi, column 19, lines 23-31] discloses "For example, if it is an n-gram language model, then the analysis would consist of identifying the individual words and word bigrams and trigrams in the hypothesis. A list of basic components in the hypotheses is generated 1806, and credit is assigned to these basic units by raising the probabilities for the basic units 1808. Then, all the basic probabilities in the language model are re-normalized 1810 which has the effect of slightly lowering all other basic probabilities." An n-gram model searches and finds the probabilities within a number of words or phrases of a target word or phrase to find the most likely candidate which would have the lowest distance. An n-gram model may be extended by any number of

n phrases or words to inspect posterior and anterior phrases or words to the target until all non-retrieved words or phrases are exhausted to get the best possible probability. The total value computed is taught by the statistical correlation value that is developed from the n-gram model, which depends on the value of n that is used.

- Horiguchi is analogous art to Gastaldo because Horiguchi both inventions deal with natural language translation. It would be obvious to someone of ordinary skill in the art at the time of the invention to combine Horiguchi with the Gastaldo device because "syntactic analysis and data structure for translation knowledge in an example-based spoken language translation are provided."
[Horiguchi, column 5, lines 4-66]

As per claim 31, Gastaldo teaches:

- a search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory
- [Gastaldo, column 5, lines 60-66] discloses "...control unit 14 queries the phrase-indexed sentence database 17 in step 41. Since all the sentences in the phrase-indexed sentence database 17 are directly indexed by noun phrases, the system is able to retrieve all pairs of sentences indexed by the selected noun phrase without any substantive delay." Gastaldo discloses that

retrieval of natural sentences corresponding to sentence pairs of original and translated sentences is performed for original sentences.

Gastaldo fails to teach,

- a parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side concordant with the interval from the object phrases to be rendered upon detection of the specific anterior words and from the anterior words extracted in the preceding search process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and

concomitantly, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side concordant with the interval from the object phrases to be rendered upon detection of the specific posterior words and from the posterior words extracted in the preceding search process till the non-retrieved posterior words have ceased to exist in the first posterior group of words

Horiguchi, in analogous art, teaches the above limitations,

- [Horiguchi, column 19, lines 23-31] discloses "For example, if it is an n-gram language model, then the analysis would consist of identifying the individual words and word bigrams and trigrams in the hypothesis. A list of basic components in the hypotheses is generated 1806, and credit is assigned to these basic units by raising the probabilities for the basic units 1808. Then, all the basic probabilities in the language model are re-normalized 1810 which has the effect of slightly lowering all other basic probabilities." An n-

gram model searches and finds the probabilities within a number of words or phrases of a target word or phrase to find the most likely candidate which would have the lowest distance. An n-gram model may be extended by any number of n phrases or words to inspect posterior and anterior phrases or words to the target until all non-retrieved words or phrases are exhausted to get the best possible probability. Horiguchi teaches the third evaluation value by disclosing an n-gram model.

- Horiguchi is analogous art to Gastaldo because Horiguchi both inventions deal with natural language translation. It would be obvious to someone of ordinary skill in the art at the time of the invention to combine Horiguchi with the Gastaldo device because "syntactic analysis and data structure for translation knowledge in an example-based spoken language translation are provided."
[Horiguchi, column 5, lines 4-66]

Claim 36 is rejected under the same grounds as Claim 30 for having parallel limitations because Claim 30's system performs the method as described in Claim 36 and has been rejected above.

Claim 37 is rejected under the same grounds as Claim 31 for having parallel limitations because Claim 31's system performs the method as described in Claim 37 and has been rejected above.

Claim 41 is rejected under the same grounds as Claim 30 for having parallel limitations because Claim 30's system performs the method for a computer program as described in Claim 41 and have been rejected above. Gastaldo discloses the use of a computer program in [Gastaldo, column 4, lines 35-38] "The interface software may for instance be written in Visual Basic, and the resulting application can be an OLE server which can be integrated directly to any Visual Basic or C/C++ code."

Claim 32 rejected under 35 U.S.C. 103(a) as being unpatentable over Gastaldo et al. (Patent # 6473729) in view of Horiguchi et al. (Patent # 6243669) and further in view of Bilenko et al. (NPL document "Learning to Combine Trained Distance Metrics for Duplicate Detection in Databases").

As per claim 32, claim 31 is incorporated and Gastaldo fails to teach:

- wherein said translation identifying means identifies at least the parallel version of the object phrases to be rendered in the parallel sentence of the natural sentence of the result with shorter intra-word distance in the computed integrated recognition as the parallel translation at least of the target phrases to be rendered in the original text, searching, regarding the natural sentence extracted by the search unit, to see whether the specific front words non-retrieved and with minimal distance from the object phrases to be rendered among the. Second anterior group of words

located on the front compared to the object phrases to be rendered in the original text, excluding the anterior words adjacent to the object phrases to be rendered are aligned on the front of the object phrases to be rendered in the natural sentence or anterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific front words when discovered and the object phrases to be rendered or the front words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding front side to the distance added to the interval computed regarding the front-side words discovered in the preceding search until the non-retrieved front words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the front side respectively, and also, searching, regarding the natural sentence extracted by the search unit, to see whether the specific back-side words non-retrieved and with minimal distance from the object phrases to be rendered among the second posterior group of words located on the back side from the object phrases to be rendered in the original text, excluding the posterior words adjacent to the object phrases to be rendered are aligned on the back of the object phrases to be rendered in the natural sentence or posterior group of words extracted in

the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific back-side words when discovered and the object phrases to be rendered or the back-side words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding back side to the distance added to the interval computed regarding the back-side words discovered in the preceding search until the non-retrieved back-side words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the back side respectively, performing a comprehensive evaluation by adding up the final recognition on the anterior and posterior sides

Horiguchi, in analogous art, teaches,

- said translation identifying means identifies at least the parallel version of the object phrases to be rendered in the parallel sentence of the natural sentence of the result with shorter intra-word distance in the computed integrated recognition as the parallel translation at least of the target phrases to be rendered in the original text, searching, regarding the natural sentence extracted by the search unit, to see whether the specific front words non-retrieved and with minimal distance from the object phrases to be rendered among the anterior and posterior groups of words

- [Horiguchi, column 19, lines 23-31] discloses "For example, if it is *an n-gram language model, then the analysis would consist of identifying the individual words and word bigrams and trigrams in the hypothesis*. A list of basic components in the hypotheses is generated 1806, and credit is assigned to these basic units by raising the probabilities for the basic units 1808. Then, all the basic probabilities in the language model are re-normalized 1810 which has the effect of slightly lowering all other basic probabilities." An n-gram model searches and finds the probabilities within a number of words or phrases of a target word or phrase to find the most likely candidate which would have the lowest distance. An n-gram model may be extended by any number of n phrases or words to inspect posterior and anterior phrases or words to the target until all non-retrieved words or phrases are exhausted to get the best possible probability. Horiguchi teaches the third evaluation value by disclosing an n-gram model.
- Horiguchi is analogous art to Gastaldo because Horiguchi both inventions deal with natural language translation. It would be obvious to someone of ordinary skill in the art at the time of the invention to combine Horiguchi with the Gastaldo device because "syntactic analysis and data structure for translation knowledge in an example-based spoken language translation are provided."
- [Horiguchi, column 5, lines 4-66]

Gastaldo and Horiguchi fail to teach,

- regarding the natural sentence extracted by the search unit, to see whether the specific front words non-retrieved and with minimal distance from the object phrases to be rendered among the second anterior group of words located on the front compared to the object phrases to be-rendered in the original text, excluding the anterior words adjacent to the object phrases to be rendered are aligned on the front of the object phrases to be rendered in the natural sentence or anterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific front words when discovered and the object phrases to be rendered or the front words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding front side to the distance added to the interval computed regarding the front- side words discovered in the preceding search until the non-retrieved front words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the front side respectively, and also, searching, regarding the natural sentence extracted by the search unit, to see whether the specific back-side words non-retrieved and with minimal distance from the object phrases to be rendered among the second posterior group of words located on the back side from the object phrases to be rendered in the

original text, excluding the posterior words adjacent to the object phrases to be rendered are aligned on the back of the object phrases to be rendered in the natural sentence or posterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific back-side words when discovered and the object phrases to be rendered or the back-side words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding back side to the distance added to the interval computed regarding the back-side words discovered in the preceding search until the non-retrieved back-side words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the back side respectively, performing a comprehensive evaluation by adding up the final recognition on the anterior and posterior sides

Bilenko, in analogous art, teaches the above limitations,

- [Bilenko, Page 2] discloses "There are several well-known methods for estimating similarity between strings, which can be roughly separated into two groups: token-based techniques and character-based techniques. Jaccard and vector-space cosine similarity are examples of difference measures that operate on tokens, treating a string as a "bag of words" [21]. Character-based

measures, such as Levenshtein distance and its variants [9], compute similarity between strings by estimating the minimum sequence of changes that transform one string into another." Bilenko teaches that there are multiple methods for determining string distances and specifically [Bilenko, Page 5] discloses forward and backward algorithms for anterior and posterior matching from a corpus. [Bilenko, Page 2] discloses "the utility of different metrics is task-dependent, and therefore it is also preferable to adaptively learn an appropriate function for combining them [2]. In our approach, Support Vector Machines (SVM's) [23], are used to learn a function of multiple similarity metrics that best discriminates duplicates from non-duplicates." Thus, Bilenko teaches that multiple methods for string distancing are used and combined to formulate the best match.

- Bilenko and Gastaldo are analogous art because both deal with distancing techniques for string analysis. It would be obvious to combine Bilenko with the Gastaldo device because "first, a set of similarity metrics are trained to appropriately determine the similarity of different field values. Next, a final predicate for detecting duplicate records is learned from multiple (trained and static) similarity metrics applied to each of the individual fields. Experimental results on real and synthetic datasets show that MARLIN is more accurate than traditional techniques." [Bilenko, Page 2]

Conclusion

9. Refer to PTO-892, Notice of References Cited for a listing of analogous art.
10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to GREG A. BORSETTI whose telephone number is (571)270-3885. The examiner can normally be reached on Monday - Thursday (8am - 5pm Eastern Time).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chameli Das can be reached on 571-272-3696. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Greg A. Borsetti/
Examiner, Art Unit 4141

/CHAMELI C. DAS/
Supervisory Patent Examiner, Art Unit 4141